Increased Efficiency of Quantum State Estimation Using *Non-Separable* Measurements

Paul B. Slater
ISBER, University of California, Santa Barbara, CA 93106-2150
e-mail: slater@itp.ucsb.edu, FAX: (805) 893-7995
(February 1, 2008)

We address the "major open problem" of evaluating how much increased efficiency in estimation is possible using *non-separable* — as opposed to separable — measurements of N copies of m-level quantum systems. First, we study the six cases $m = 2$, $N = 2, \ldots, 7$ by computing the 3×3 Fisher information matrices for the corresponding *optimal* measurements recently devised by Vidal *et al* (Phys. Rev. A 60, 126 [1999]). We obtain simple polynomial expressions for the ("Gill-Massar") traces of the products of the inverse of the quantum Helstrom information matrix and these Fisher information matrices. The six traces *all* have *minima* of $2N - 1$ in the *pure state* limit — while for *separable* measurements (Phys. Rev. A 61, 042312 [2000]), the traces can equal N, but *not* exceed it. Then, the result of an analysis for $m = 3$, $N = 2$ leads us to *conjecture* that for optimal measurements for *all* m and N, the Gill-Massar trace achieves a *minimum* of $(2N - 1)(m - 1)$ in the *pure state* limit.

PACS Numbers 03.67.-a, 89.70.+c, 02.50.-r

Contents

I Introduction 2

II Proportionality between Helstrom and Fisher Information Matrices 3

III Analyses of *Optimal* Measurements of Vidal *et al* for N Copies of Two-Level Quantum Systems 4
- A Computation of the Fisher Information Matrices . . . 4
  - 1 $N = 2$ . . . 4
  - 2 $N = 3$ . . . 4
  - 3 $N = 4$ . . . 4
  - 4 $N = 5$ . . . 5
  - 5 $N = 6$ . . . 5
  - 6 $N = 7$ . . . 5
  - 7 $N > 7$ . . . 6
- B Properties of the Computed Fisher Information Matrices . . . 6
  - 1 *Diagonal* nature for *even* N in spherical coordinates . . . 6
  - 2 Pure- and fully mixed state limits . . . 6
  - 3 Integrals over Bloch sphere of volume elements . . . 7
- C Gill-Massar Traces . . . 7
- D Analyses for m-Level *Pure* States . . . 8
  - 1 $m = 2$ . . . 8
  - 2 $m = 3$ . . . 8
  - 3 supplementary analysis for 3-level *mixed* states . . . 9
  - 4 $m = 4$ . . . 9

IV Universal Coding 9
- A Classical results of Clarke and Barron . . . 10
- B Quantum Results of Krattenthaler and Slater for Two-Level Systems . . . 10
- C Relations between *Monotone Metrics* and the Fisher Information Matrices Computed in Sec. III A 12
  - 1 The (2,2)- and (3,3)-entries of the diagonal Fisher information matrices for even N . . . 12
  - 2 The (1,1)-entries of the diagonal Fisher information matrices for even N . . . 13
  - 3 *Modified* Gill-Massar traces based on the Yuen-Lax (maximal monotone) and quasi-Bures information matrices . . . 13

I. INTRODUCTION

We investigate information-theoretic properties of the optimal measurement schemes recently devised by Vidal *et al* [1], helping thereby to address the "major open problem" [2] of evaluating how much increased efficiency in estimation is possible using *non-separable* measurements (cf. [3]). In their extensive study, "State estimation for large ensembles," which we seek to extend here, Gill and Massar stated that "we cannot compare our results with the recent analysis of covariant [optimal] measurements on mixed states [1] because we suppose separability of the measurement, whereas [1] does not" [2]. A "separable measurement is one that can be carried out sequentially on separate particles, where the measurement on one particle at any stage (and indeed which particle to measure: one is allowed to measure particles several times) can depend arbitrarily on the outcomes so far" [2].

The analyses here are conducted in terms of the (classical) *Fisher information* (of the probability distributions associated with the non-separable measurements), making use of the quantum (Helstrom) Cramér-Rao bound [4] on the Fisher information matrix for any *oprom* (operator-valued probability measure) [5,6]. Contrastingly, the studies of Vidal and his several Barcelona colleagues [1,7–9] have been formulated primarily in terms of *fidelity*, $F(\rho, \rho')$ (ρ and ρ' being density matrices) [10,11], and secondarily, *information gain* [7]. Now, there surely exists an intimate connection between these approaches, since $2(1 - F(\rho, \rho'))$ functions as the *Bures* distance between ρ and ρ'. The Bures metric is a distinguished member (the *minimal* one) of a continuum of possible quantum extensions — each associated with a distinct *operator monotone* function — of the (classical) Fisher information metric [12–14]. The Helstrom-Cramér-Rao bound corresponds to the particular use of the Bures metric *via* the concept of the *symmetric logarithmic derivative* [4]. An interesting hypothesis is that asymptotically the Fisher information matrix for optimal measurements is simply proportional to the metric tensor associated with some specific operator monotone function. (Our results below indicate that such a role is definitely *not* played by the Bures metric.)

We shall be concerned here primarily (cf. secs. III D 2 and III D 4) with the two-level quantum systems, representable by the 2×2 density matrices,

$$\rho = \frac{1}{2} \begin{pmatrix} 1+z & x+\mathrm{i}y \\ x-\mathrm{i}y & 1-z \end{pmatrix}, \tag{1}$$

where $r^2 = x^2 + y^2 + z^2 \leq 1$. The particular (x, y, z) parameterization employed in (1) corresponds to the use of Cartesian coordinates for the "Bloch (or Poincaré) sphere" (unit ball in three-space) representation of the two-level systems [15] [16, sec. 4.2], while the alternative (spherical coordinate) parameter r is the radial distance from the origin. Pure states, for which $|\rho| = 0$, correspond to $r = 1$ and the fully mixed state, for which $|\rho| = \frac{1}{4}$, to $r = 0$.

For the cases of N copies ($N = 2, \ldots, 7$) of a two-level quantum system (1) we obtain below in sec. III C a quite interesting pattern of results of increased efficiency using non-separable measurements, which strongly suggests generalizability to arbitrary N. To explicitly examine the cases $N > 7$ would either entail considerable additional computations for each specific N and/or substantial analytical advances (cf. sec. IV C) allowing one to formally establish the measure of increased efficiency for *arbitrary* N. (We note that Latorre *et al* [8] had to proceed *case-by-case*, that is, each N individually, since they "did not know how to build the POVM algorithmically".) In sec. IV C we explore one possible approach in this regard, attempting to explain the Fisher information matrices we compute in sec. III in terms of monotone metrics. In sec. III, we also formulate a conjecture as to the increase in efficiency achieveable using non-separable optimal measurements for N copies of m-level quantum systems in general.

To begin our study, immediately below in sec. II, we expand upon an observation [17, p. 2684] regarding an information-theoretic relationship between certain classical and quantum entities — that is, the Fisher information matrix for a certain (quadrinomial) multinomial probability distribution and the quantum Helstrom information matrix (proportional to the Bures metric tensor), and its implications for optimal measurements.

In sec. IV we examine further ramifications on issues of state estimation [2,4] and universal coding (data compression) [18–21]. There appears to be an interesting relation between the devising of optimal measurements as in [1], and universal quantum coding, as both processes involve averaging with respect to isotropic prior probability distributions by "projecting onto total spin eigenspaces, and within each such subspace, onto total spin eigenstates with maximal total spin component in some direction" [1] — cf. [1, eqs. (5.33) and (5.34)] and [19, eq. (2.48)]. The particular prior distribution which yields both the minimax and maximin for the universal quantum coding of the two-level systems is based on the *quasi-Bures* metric, a particular example of a monotone metric. We attempt in sec. IV C to relate the Fisher information matrices we compute in sec. III to the monotone metrics.

II. PROPORTIONALITY BETWEEN HELSTROM AND FISHER INFORMATION MATRICES

The density matrices (1) turn out to have an intimate relationship with a particular form of multinomial (that is, quadrinomial) probability distributions — the *four* distinct possible outcomes being assigned probabilities

$$x^2, \quad y^2, \quad z^2, \quad 1-x^2-y^2-z^2. \tag{2}$$

One can attach to the three-dimensional convex set of two-level quantum systems (1), adapting one (the simplest) of the "explicit" formulas of Dittmann [22, eq. (3.7)] [23],

$$d_{Bures}(\rho, \rho + \mathrm{d}\rho)^2 = \frac{1}{4}\mathrm{Tr}\{\mathrm{d}\rho \mathrm{d}\rho + \frac{1}{|\rho|}(\mathrm{d}\rho - \rho \mathrm{d}\rho)(\mathrm{d}\rho - \rho \mathrm{d}\rho)\}, \tag{3}$$

the 3×3 quantum (Helstrom) information matrix [4,2,24] (that is, *four* times the Bures metric tensor [23,25,26,13]),

$$H_q(x,y,z) = \frac{1}{(1-x^2-y^2-z^2)} \begin{pmatrix} 1-y^2-z^2 & xy & xz \\ xy & 1-x^2-z^2 & yz \\ xz & yz & 1-x^2-y^2 \end{pmatrix}. \tag{4}$$

We use the subscripts q and c — in a suggestive, perhaps not fully rigorous manner — to denote results stemming from quantum or classical considerations. Also, note that (4) "blows up" at the pure states themselves — so it will be problematical, at best, to directly compare results pertaining to (4) with ones based on *pure state* models [2,27].

In spherical coordinates (r, θ, ϕ), $x = r\cos\theta, y = r \sin\theta \cos\phi, z = r\sin\theta\sin\phi$, the matrix (4) takes a *diagonal* form,

$$H_q(r,\theta,\phi) = \begin{pmatrix} \frac{1}{1-r^2} & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2\sin^2\theta \end{pmatrix}, \tag{5}$$

for this *orthogonal* system of coordinates (cf. [28]). (Below, in the interest of succinctness, we will replace the frequently-occurring expression $x^2+y^2+z^2$ by its equivalent, r^2.)

Now, the quantum information matrices (4) and (5) are simply proportional to the (classical) Fisher information [29] matrices $I_c(x,y,z)$ and $I_c(r,\theta,\phi)$ for the quadrinomial probability distribution (2). (By way of algorithmic example, the xy-entry of the 3×3 Fisher information matrix — in its Cartesian coordinate form, $I_c(x,y,z)$ — is computable as the expected value of the [two-fold] product of the logarithmic derivatives of (2) with respect to x and with respect to y.) More precisely, the nine entries of $I_c(x,y,z)$ are all *four* times the corresponding entries of (4), that is

$$I_c(x,y,z) = 4H_q(x,y,z). \tag{6}$$

A natural explanation for this phenomenon is that the *information geometry* [30] of both models is that of the standard metric on the surface of a three-sphere in four-dimensional Euclidean space [13,31].

Both quantum (Helstrom) information and Fisher information possess the property of *additivity*, that is, for N independent identical density matrices or probability distributions, the information matrices (possibly scalars) are N times those for a single one [5, exer. 1.10] [4, sec. VI.4] [32–35].

By the quantum version of the Cramér-Rao theorem [4], the inverse matrix $H_q(x,y,z)^{-1}$ serves as a lower bound on the variance-covariance matrix $V(x,y,z)$ for any *unbiased* estimator of the parameters (x,y,z) of ρ. (This means that the matrix difference, $V(x,y,z) - H_q(x,y,z)^{-1}$, must be nonnegative definite, that is, have all its eigenvalues nonnegative.) In this regard,

$$H_q(x,y,z)^{-1} = \begin{pmatrix} 1-x^2 & -xy & -xz \\ -xy & 1-y^2 & -yz \\ -xz & -yz & 1-z^2 \end{pmatrix} \tag{7}$$

(Of course, $H_q(r,\theta,\phi)^{-1}$ is diagonal.)

By dint of the additivity of information, in conjunction with the Cramér-Rao theorem (cf. [2, eq. (26)]), one can conclude that it is *not* possible to devise for $N < 4$ independent identical two-level systems, an *oprom* [5,6], which has for its outcomes the quadrinomial distribution (2) (cf. [1,36]). (When we attempted to construct such an oprom for the case $N = 2$, we found that the four operators could *not* all be nonnegative definite if they were to yield (2).) However, for $N \geq 4$, the question of whether such an oprom exists would appear to be a completely open one — since now the Cramér-Rao theorem does *not* rule out its possibility. (The results of Vidal *et al* [1] show that an optimal *minimal* number of measurements for $N > 3$ is at least *fifteen*, exceeding the number *four* for an oprom that would give as its outcomes, the quadrinomial probability distribution (2).) If such an oprom could be found for $N = 4$ itself, then the Cramér-Rao inequality would be *fully* saturated.

III. ANALYSES OF *OPTIMAL* MEASUREMENTS OF VIDAL *ET AL* FOR N COPIES OF TWO-LEVEL QUANTUM SYSTEMS

A. Computation of the Fisher Information Matrices

1. $N = 2$

Let us now consider the probability distribution in [1] obtained from the optimal minimal number (five) of measurements for the case of $N = 2$ identical independent copies of the two-level systems (1). The five probabilities — as we have explicitly found — can be written as (the three)

$$\frac{1}{4}(1-r^2), \quad \frac{3}{16}(1+z)^2, \quad \frac{1}{48}(8x^2 - 4\sqrt{2}x(z-3) + (z-3)^2), \tag{8}$$

together with the pair

$$\frac{1}{48}(9 + 2x^2 \pm 4\sqrt{3}xy + 6y^2 + 2\sqrt{2}(x \pm \sqrt{3}y)(z-3) - 6z + z^2).$$

Quite remarkably, the associated Fisher information matrix ($\tilde{I}_c$) turns out to precisely equal the quantum (Helstrom) information matrix, $H_q(x,y,z)$ — and not $2H_q(x,y,z)$, which is the upper bound furnished by the quantum Cramér-Rao theorem. So, the bound could be said to be "half-saturated". (In regard to this specific result, R. Gill has observed that there may exist other measurement schemes which are *sub-optimal* accoding to the *fidelity* criterion of [1], but superior in terms of Fisher information (cf. [7]).)

2. $N = 3$

For an optimal minimal set of measurements for $N = 3$, we can take the eight probabilities, consisting of the four pairs,

$$\frac{(1 \pm x)^3}{12}, \quad \frac{(1 \pm y)^3}{12}, \quad \frac{(1 \pm z)^3}{12}, \quad \frac{1}{4}(1 \pm \frac{x+y+z}{\sqrt{3}})(1-r^2). \tag{9}$$

The associated Fisher information matrix is expressible as

$$2H_q(x,y,z) + \frac{1}{2((x+y+z)^2-3)} \begin{pmatrix} a & b & b \\ b & a & b \\ b & b & a \end{pmatrix}, \tag{10}$$

where $a = 2(1 - xy - xz - yz)$ and $b = -1 + r^2$. The second summand in (10) is *negative* definite (having two of its three negative eigenvalues equal to $-\frac{1}{2}$), while $3H_q(x,y,z)$ is the upper bound on the Fisher information matrix provided by the Cramér-Rao theorem.

3. $N = 4$

An optimal minimal set of measurements for $N = 4$ yields a fifteen-vector of probabilities. The Fisher information matrix for this probability distribution is

$$3H_q(x,y,z) + \frac{1}{12} \begin{pmatrix} -7-5y^2-5z^2 & 5xy & 5xz \\ 5xy & -7-5x^2-5z^2 & 5yz \\ 5xz & 5yz & -7-5x^2-5y^2 \end{pmatrix}. \tag{11}$$

The second term is *negative* definite with one eigenvalue equal to $-\frac{7}{12}$ and the other two, $-\frac{1}{12}(7+5r^2)$. If we subtract (11) from the Cramér-Rao upper bound $4H_q(x,y,z)$, we obtain (as we must) a nonnegative definite matrix, having two eigenvalues $\frac{1}{12}(19+5r^2)$ and one, $\frac{7}{12} + \frac{1}{1-r^2}$.

4. $N=5$

For $N=5$, a twenty-vector of probabilities was obtained for the optimal minimal number of measurements. The Fisher information matrix can be expressed as the sum of $4H_q(x,y,z)$ (which dominates it, while $3H_q(x,y,z)$ does not) and a *negative* definite matrix, having one of its three negative eigenvalues equal to $-\frac{3}{16}(5+3r^2)$. This negative definite matrix can be written as the product of $\frac{1}{16(-3+(x+y+z)^2)}$ and a 3×3 matrix, the $(1,1)$ cell of which is

$$-2(-20+7y^4+9y^3z-11z^2+7z^4-5x^3(y+z)+3yz(5+3z^2)+ \tag{12}$$

$$3x(y+z)(5+3y^2+3z^2)+x^2(10+7y^2-5yz+7z^2)+y^2(-11+14z^2))$$

and the $(1,2)$ off-diagonal entry is

$$-5x^4+14x^3y+2x^2(5+9y^2+14yz-5z^2)-5(-1+y^2+z^2)^2+14xy(-3+(y+z)^2). \tag{13}$$

The remaining cells are obtainable by simple symmetry arguments (for example, the (2,2) cell can be gotten by interchanging x and y in (12)).

5. $N=6$

For $N=6$, we used an optimal (but not minimal) set of thirty-three measurements. We found — using a large number of randomly generated points (x,y,z) — that the associated Fisher information matrix was strictly dominated by $5H_q(x,y,z)$, but not by $4.99H_q(x,y,z)$. The Fisher information matrix takes the form (cf. (11))

$$5H_q(x,y,z)+\frac{1}{120}\begin{pmatrix} a & Axy & Axz \\ Axy & b & Ayz \\ Axz & Ayz & c \end{pmatrix}, \tag{14}$$

where

$$A=193-31r^2, \quad a=-125-146y^2-146z^2+31(y^2+z^2)^2+x^2(47+31y^2+31z^2), \tag{15}$$

and the diagonal entry b can be obtained from a by interchanging x and y, and c from a by interchanging x and z.

One of the three negative eigenvalues of the second ("residual") matrix in (14) is $(125-172r^2+47r^4)/(120(-1+r^2))$. Now, if we were to rewrite (14) in the form of $4.99H_q(x,y,z)$ plus a *slightly* revised residual matrix, the eigenvalue in question would be altered only in the respect that the constant 125 would change to 123.8. This would render it *positive* for $r>.992348$, leading to a loss of strict dominance for $r\in[.992348,1]$. In this specific sense, the upper bound of $5H_q(x,y,z)$ on the Fisher information matrix is *tight*. The residual matrix for $N=4$ strictly dominates that for $N=6$. This indicates that the "fit" of $(N-1)H_q(x,y,z)$ to the Fisher information matrix for optimal measurements of N copies *improves* as N increases.

6. $N=7$

For $N=7$, employing a 42-vector of probabilities, we found the Fisher information matrix to be strictly dominated by $6H_q(x,y,z)$, but *not* by $5.99H_q(x,y,z)$. Reviewing our previous analyses, we then found that the analogous situation held also for $N=3,\ldots,6$, that is, the Fisher information matrix was dominated by $(N-1)H_q(x,y,z)$, but not by $(N-1.01)H_q(x,y,z)$. The violations of these *diminished* bounds occur for nearly pure states, that is $r\approx 1$.

Pursuing this line of thought, if we restrict consideration to the more mixed states for which $r<\frac{1}{2}$, then for $N=7$ we have found that $3.9H_q(x,y,z)$, but not $3.85H_q(x,y,z)$ bounds the Fisher information matrix for the optimal set of measurements. Calculations suggest the hypothesis that in the neighborhood of the fully mixed state $r=0$, the bound on the Fisher information matrices approaches from above $NH_q(0,0,0)/2$, that is $\frac{N}{2}$ times the 3×3 identity matrix. Now, the fully mixed state is classical (binomial) in character, while the pure states are quantum in nature. (It is interesting to note that Frieden finds that in classical scenarios, only *one-half* of the bound or phenomenological information J is utilized in the intrinsic quantum information I [29, eqs. (5.39), (6.55)]. "In all covariant quantum theories (e. g., quantum mechanics, quantum gravity) I and J are exactly equal. In deterministic classical theories such as classical electromagnetics and general relativity $I=J/2$. But in statistical classical theories $I=J$ again" [e-mail message from Frieden].)

7. $N > 7$

We are not able to proceed any further, that is for $N > 7$, as there presently do not appear to be corresponding sets of optimal measurements. As a *caveat* to the reader, let us point out that to recreate the optimal measurements for the cases $N = 6$ and 7 (which unlike the instances $N < 6$, were not formally demonstrated to be minimal in character), it is necessary to rely upon the quant-ph preprint version (9803066) of [8], since there are certain errors (as confirmed in an e-mail from R. Tarrach, though no formal *erratum* has appeared) in the final, published paper.

B. Properties of the Computed Fisher Information Matrices

1. Diagonal nature for even N in spherical coordinates

We have found that the Fisher information matrices given above for the optimal measuements of Vidal *et al* [1] for both $N = 4$ and 6 are *diagonal* in spherical coordinates (r, θ, ϕ). For $N = 4$, this is

$$\frac{1}{12}\begin{pmatrix} \frac{29+7r^2}{1-r^2} & 0 & 0 \\ 0 & r^2(29-5r^2) & 0 \\ 0 & 0 & r^2(29-5r^2)\sin^2\theta \end{pmatrix}, \tag{16}$$

and for $N = 6$,

$$\frac{1}{120}\begin{pmatrix} \frac{475+172r^2-47r^4}{1-r^2} & 0 & 0 \\ 0 & r^2(475-146r^2+31r^4) & 0 \\ 0 & 0 & r^2(475-146r^2+31r^4)\sin^2\theta \end{pmatrix}. \tag{17}$$

For $N = 2$, we also have a corresponding diagonal matrix, that is, (5).

Cox and Reid [37, p. 2] have listed three "consequences of orthogonality" of the parameterization of a Fisher information matrix, such as we have just observed. These are that: (i) the maximum likelihood estimates of the means of the parameters are asymptotically independent; (ii) the asymptotic standard error for estimating one parameter is the same whether the other parameters are treated as known and unknown; and (iii) there may be simplifications in the numerical determination of the means of the parameters. "While orthogonality can always be achieved locally, global orthogonality is possible only in special cases" [37, p. 2]. In accompanying discussions to [37], Sweeting identifies four advantages to orthogonalization — computation, approximation, interpretation, and elimination of nuisance parameters — while Barndorff-Nielsen, as well as Moolgavkar and Prentice, explain parameter orthogonality in terms of Frobenius' Theorem. The latter authors also indicate that the theorem of de Rham [38, p. 187] gives necessary and sufficient conditions for each orthogonal parameter to be independent of the others (as they are *not* in our three even-dimensional examples just given).

2. Pure- and fully mixed state limits

Again using spherical coordinates, it is interesting to note that for the *odd* cases of $N = 3, 5, 7$, in the pure state limit ($r \to 1$), the off-diagonal elements of the corresponding 3×3 Fisher information matrix converge to zero. In all six (both odd and even) cases, in this same limit, the (1,1)-entries are indeterminate, the (2,2)-entries are $\frac{N}{2}$ and the (3,3)-entries are $\frac{N \sin^2\theta}{2}$.

For the fully mixed state, $r = 0$ (allowing the angular variables θ and ϕ to remain free), the only non-zero entry is the (1,1)-cell. For $N = 2$ it is 1, for $N = 3$ it is

$$\frac{1}{6}(10 + \sin 2\theta(\cos\phi + \sin\phi) + \sin^2\theta \sin 2\phi), \tag{18}$$

for $N = 4$ it is $\frac{29}{12}$, for $N = 5$, it is $\frac{(103+5\cos 2\phi)}{32}$, for $N = 6$ it is $\frac{95}{24}$, and for $N = 7$,

$$\frac{1}{96}(456\cos^2\theta + 7\sin 2\theta(\cos\phi + \sin\phi) + \sin^2\theta(456 + 7\sin 2\phi)). \tag{19}$$

3. *Integrals over Bloch sphere of volume elements*

For $N = 2$, the integral of the volume element of the Fisher information matrix (that is, the square root of the determinant) over the (Bloch sphere of) two-level quantum systems is $\pi^2 \approx 9.8696$, for $N = 3$ it is 21.0235, for $N = 4$, it is

$$\frac{1}{441}\sqrt{\frac{29}{3}}\pi(4705E(-\frac{7}{29}) - 4194K(-\frac{7}{29})) \approx 35.0281 \tag{20}$$

(where E and K denote the corresponding elliptic integrals), for $N = 5$, it is 51.0763, for $N = 6$, it is 69.1253, and for $N = 7$, 88.8621. These particular results would be needed for the application to the optimal measurements of Vidal *et al* [1] of the universal coding theorem of Clarke and Barron [18], discussed below in sec. IV A.

C. Gill-Massar Traces

Let us first observe that Gill and Massar [2, eq.(26)] asserted that the upper (quantum [Helstrom] Cramér-Rao) bound NH_q, was *not*, in general, achievable in a multiparameter setting. This does appear to be strictly the case. However, our results for $N = 2, \ldots, 7$ for the three-parameter 2×2 density matrices, indicate that — using the optimal measurements of Vidal *et al* [1] — one can, by choosing N large enough, come indefinitely close for the nearly pure states to this bound.

To further relate to these analyses of Gill and Massar, we have computed for $N = 2, \ldots, 7$, the traces of the product of $H_q(x, y, z)^{-1}$, given in (7), and the Fisher information matrices we have obtained using the optimal measurements of Vidal *et al*. (The traces of Fisher information matrices play a central role in the work of Frieden on the fundamental equations of physics [29, sec. 2.3.2].) For the estimation of pure states, Theorem I in [2] asserts that this trace quantity is bounded above by N, while Theorem II there says that the same bound applies to mixed states, with the restriction to *separable* measurements. It is also demonstrated there that these bounds are attainable — and for large N *simultaneously* for *all* states.

For $N = 2$, it is easy to see, in the context of the results above, that this ("Gill-Massar") trace result is simply 3. For $N = 3$, we get another constant, 5, for the trace. For $N = 4$, we obtain

$$GM_4 = \frac{29 - r^2}{4}, \tag{21}$$

which is 7 for pure states and 7.25 for the fully mixed state. For $N = 5$, the Gill-Massar trace is

$$GM_5 = \frac{19 - r^2}{2}, \tag{22}$$

which is 9 for pure states and 9.5 for the fully mixed state. For $N = 6$, it is

$$GM_6 = \frac{95 - 8r^2 + r^4}{8}. \tag{23}$$

This last expression is monotonically decreasing from $\frac{95}{8} = 11.875$ at $r = 0$ to 11, that is, $2N - 1$ at $r = 1$. For $N = 7$, the Gill-Massar trace is

$$GM_7 = \frac{57 - 6r^2 + r^4}{4}, \tag{24}$$

which equals $\frac{57}{4} = 14.25$ at $r = 0$ and 13 at $r = 1$, being again $2N - 1$. (In an earlier version of this paper, quant-ph/0002063, the results given — including Fig. 1, plotting the Gill-Massar trace — for $N = 7$ were "anomalous", in this regard. We subsequently ascertained that they were erroneous in nature, due to a programming error.) In Fig. 1, we plot $\frac{GM_N}{(2N-1)}$ for $N = 4, 5, 6$ and 7.

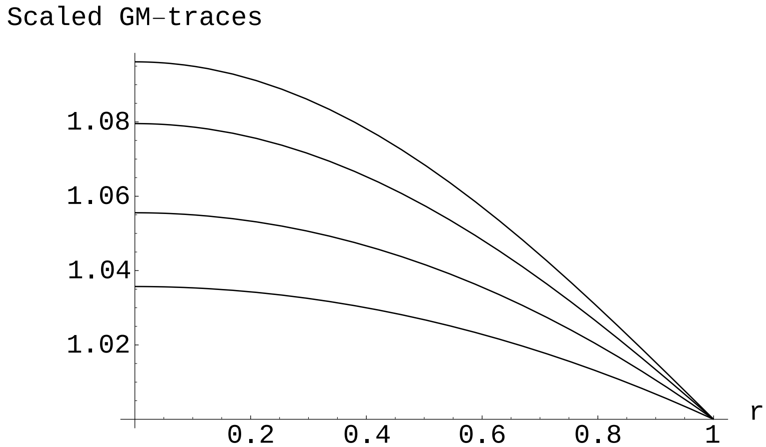

FIG. 1. Gill-Massar traces for $N = 4, 5, 6$ and 7 scaled by their values at the pure states, $r = 1$, that is, $2N - 1$. The y-intercepts for $r = 0$, corresponding to the fully mixed state, increase with N.

It is easy to see, then, that in these six cases the Gill-Massar bound [2, eq. (27)] of N is violated — as Theorem III of their paper recognizes will occur for *non-separable* measurements. So, we obtain a simple pattern of $2N - 1$ for the minimum of the trace quantity in question. In regards to these results, R. Gill remarked in an e-mail message of Feb. 18, 2000 that "this is all very interesting. It means that there is a big discontinuity at the surface of the Bloch sphere (where none of these 3×3 Fisher information matrices is well-defined), and it means that the gain in using joint measurements over separate measurements for mixed states is substantial throughout the Bloch sphere".

D. Analyses for m-Level *Pure* States

1. $m = 2$

In a further effort to relate to the analyses of Gill and Massar [2], let us consider for the moment simply the two-level pure states, so we set $r = 1$. In terms of the polar coordinates (θ, ϕ), the Helstrom information matrix takes the form (cf. (5), [40, p. 4238])

$$\begin{pmatrix} 1 & 0 \\ 0 & \sin^2\theta \end{pmatrix}. \tag{25}$$

Then, the Fisher information matrix for the optimal measurements of N copies [8] is simply $\frac{N}{2}$ times (25), as we have confirmed through computations for $N = 2, \ldots, 7$ (cf. [2]). (So, in the pure state case, unlike the mixed state one, the quantum Cramér-Rao bound of N times (25) is not asymptotically approached — though the Gill-Massar trace bound of N is achievable.)

2. $m = 3$

We have also verfied that the same basic additive relation holds in the case of the *three*-level pure states for $N = 2$, using the formulas in [9]. Let us use the parameterization of these states in terms of *four* angular variables $(\theta, \phi, \chi_1, \chi_2)$ employed in [41, eq. (2.1)],

$$|\psi\rangle = e^{i\chi_1} \sin\theta \cos\phi |1\rangle + e^{i\chi_2} \sin\theta \sin\phi |2\rangle + \cos\theta |3\rangle. \tag{26}$$

Then, the Helstrom information matrix is

$$\begin{pmatrix} 4 & 0 & 0 & 0 \\ 0 & 4\sin^2\theta & 0 & 0 \\ 0 & 0 & a & -\sin^4\theta \sin^2 2\phi \\ 0 & 0 & -\sin^4\theta \sin^2 2\phi & b \end{pmatrix}, \tag{27}$$

where (cf. [42])

$$a = \frac{1}{2}(6 + 2\cos 2\theta + \cos 2(\theta - \phi) - 2\cos 2\phi + \cos 2(\theta + \phi)) \sin^2\theta \cos^2\phi, \tag{28}$$

$$b = -\frac{1}{2}(-6 - 2\cos\theta + \cos 2(\theta - \phi) - 2\cos 2\phi + \cos 2(\theta + \phi)) \sin^2\theta \sin^2\phi.$$

(Note that (27) is free of the variables, χ_1 and χ_2 — as (5) is free of ϕ.) So, for $N = 2$ copies of a spin-1 system, the Fisher information matrix is identically (27), paralleling the specific results for both the pure and mixed two-level quantum systems for $N = 2$. We also intend to analyze the case $N = 3$, using the specific prescription for the corresponding optimal measurements in [9, sec. 6].

3. supplementary analysis for 3-level mixed states

We have attempted — following the general methodology laid out by Vidal *et al* [1] for the *two*-level mixed quantum systems — to construct an optimal measurement scheme for $N = 2$ copies of mixed *three*-level systems. In doing so, we incorporated the optimal measurements for $N = 2$ copies of *pure* three-level quantum systems presented by Acín, Latorre and Pascual in [9, sec. 5], that were utilized immediately above. (J. Latorre informs me that he and his co-authors "did not find any manageable way to make progress" in such extended $m = 3$ *mixed* cases, although he did point out that Arvind had recast and further developed many of their results using Penrose rays — in apparently yet unpublished work.) This led us to an oprom with *twelve* distinct outcomes, *nine* corresponding to the vectors explicitly presented in [9, eqs. (39), (40)], and the additional *three* coming from our own orthogonal decomposition of the associated rank three "residual" projector (cf. [1, eq. (3.3)]). (A weight of $\frac{2}{3}$ was applied to the subset of nine outcomes.)

With this twelve-outcome oprom in hand, we found by *numerical* means that the Gill-Massar trace equalled a constant, 6 (while for $N = 2$ copies of *two*-level systems this trace quantity was found in sec. III C also to be a constant, 3). (In [42], we have been investigating the possibility of *symbolically* inverting the 8×8 Helstrom information matrix — making use of a recently-developed Euler angle parameterization of the 3×3 density matrices [43]. The Gill-Massar trace would, of course, be the trace of the product of this inverse matrix and the Fisher information matrix associated with the twelve-outcome oprom.) This result and our earlier ones for $m = 2$, $N = 2, \ldots, 7$, lead us to conjecture that for non-separable optimal measurements of N m-level quantum systems, the Gill-Massar trace for all m and N is exactly $(2N - 1)(m - 1)$ in the pure state limit, and no less than this for any mixed state.

Now, for any measurement of a strictly pure state itself, the Gill-Massar trace can not exceed $N(m-1)$ by Theorem I of [2]. (This bound is known to be achieveable for $m = 2$ by Theorem VII of [2], and for mixed states using separable measurements by Theorem VI.) So there is a clear discontinuity displayed by *non-separable optimal* measurements *near* the pure state boundary, as well as considerable increased efficiency in estimating strictly mixed or impure states through the use of such measurements.

4. $m = 4$

We have ascertained the Helstrom information matrix for pure states of *four*-level systems, making use of the appropriate analogue of the parameterization (26) presented in [44, eq. (13)]. The six parameters naturally divide into two sets of three, and once again the entries of the Helstrom information matrix are free of the (three) members of one of the two sets.

IV. UNIVERSAL CODING

We can also apply to the three-dimensional family of quadrinomial probability distributions (2) certain important (classical) asymptotic results of Clarke and Barron [18] pertaining to a number of problems, including those of universal data compression and density estimation. Then, we can compare their formulas with those for the 2×2 density matrices (1), based on the extension to the quantum domain of two-level systems by Krattenthaler and Slater [19,20] of this work of Clarke and Barron (cf. [11]). (In what follows, we will denote probability distributions of a general nature by w and more specific ones by W, and subscript them — as noted before — by either c or q to denote a result stemming from an analysis in the classical or quantum domain.)

A. Classical results of Clarke and Barron

Clarke and Barron examined the relative entropy ($N \rightarrow \infty$) between a true density function and a joint ("Bayesian") density function for a sequence of N random variables taken to be the average of the possible densities (comprising a parameterized family) with respect to a (prior) probability distribution over this family of density functions. The result of Clarke and Barron for the asymptotic relative entropy (Kullback-Leibler index) between the true density and the mixture is

$$\frac{d}{2}\log\frac{N}{2\pi \mathrm{e}}+\frac{1}{2}\log|I_c(\alpha)|-\log w_c(\alpha)+o(1), \tag{29}$$

where α denotes the d-vector of variables parameterizing the family of densities, $w_c(\alpha)$ a prior probability distribution used to average the N-fold products of independent identical density functions, and $I_c(\alpha)$ the associated $d \times d$ Fisher information matrix. As applied to our particular three-parameter ($d = 3$) family of quadrinomial distributions (2), with $\alpha = (r, \theta, \phi)$, we have

$$|I_c(r,\theta,\phi)| = (\frac{64}{1-r^2})r^4\sin^2\theta. \tag{30}$$

Then, if we choose for the probability distribution, $w_c(\alpha)$, the particular one

$$W_c(r,\theta,\phi) = (\frac{1}{\pi^2\sqrt{1-r^2}})r^2\sin\theta \quad \propto \sqrt{|I_c(r,\theta,\phi)|}, \tag{31}$$

the asymptotic relative entropy between the true density and its Bayesian (mixture) average assumes the form [18, eq. (1.4)]

$$\frac{3}{2}\log\frac{N}{2\pi \mathrm{e}}+\log 8\pi^2+o(1). \tag{32}$$

(Let us note that $r^2 \sin\theta \mathrm{d}r\mathrm{d}\theta\mathrm{d}\phi$ is the Jacobian determinant of the transformation from Cartesian to spherical coordinates or, equivalently, the volume element in spherical coordinates.) Our particular selection of $W_c(r, \theta, \phi)$ is "Jeffreys' prior" for this case, that is the normalized (over the Bloch sphere) form of the volume element ($\sqrt{|I_c(r,\theta,\phi)|}$) of the Fisher information metric (cf. sec. III B 3). (The normalization factor, $8\pi^2$, is evident in (32)). Jeffreys' priors, as shown by Clarke and Barron [18], fulfill the desideratum of yielding the common *minimax* and *maximin* of the asymptotic relative entropy. In the quantum analogue, though, (31) does not play this distinguished role, although a close ("quasi-Bures") relative of it does [20,45]. This probability distribution is

$$W_q(r,\theta,\phi) = .0832258\frac{\mathrm{e}}{1-r^2}(\frac{1-r}{1+r})^{\frac{1}{2r}}r^2\sin\theta. \tag{33}$$

B. Quantum Results of Krattenthaler and Slater for Two-Level Systems

Krattenthaler and Slater [19,20] have sought to extend the general results of Clarke and Barron to the two-level *quantum* systems (1). They averaged the N-fold *tensor* products of identical 2×2 density matrices (1) (rather than averaging the simple products of N *random variables*) with respect to (spherically-symmetric/unitarily-invariant) probability distributions distributions of the form $w_q(r)r^2\sin\theta$ (cf. [1, eq. (1.4)]). The analogue (in terms of the *quantum* relative [von Neumann] entropy) of the Clarke-Barron result (29) is then ($d = 3$)

$$\frac{3}{2}\log\frac{N}{2\pi \mathrm{e}}+\frac{1}{2}\log I_q(r)-\log w_q(r)+o(1), \tag{34}$$

where (cf. (30))

$$I_q(r) = \frac{\mathrm{e}^2}{(1-r^2)^2}(\frac{1-r}{1+r})^{\frac{1}{r}}. \tag{35}$$

So,

$$I_q(r) r^4 \sin^2\theta = 144.372 W_q(r,\theta,\phi)^2, \tag{36}$$

which can be compared with its classical counterpart,

$$|I_c(r,\theta,\phi)| = 64\pi^4 W_c(r,\theta,\phi)^2, \tag{37}$$

where $64\pi^4 \approx 6234.18$.

As noted [20], the quasi-Bures probability distribution, $W_q(r,\theta,\phi)$, given by (33), fulfills in the quantum domain of two-level systems (1), the distinguished role — in yielding the common asymptotic minimax and maximin — of the Jeffreys' prior (that is, the volume element of the Fisher information metric) in the classical sector. In Fig. 2 we plot the term $\frac{1}{2}\log I_q(r)$, present in (34), along with the comparable (but always larger for $r < 1$) classical term, $\frac{1}{2}\log\frac{64}{1-r^2}$, in (30). The units of the vertical axis are, then, "nats" of information. (A nat is equal to $1/\log_e 2 \approx 1.4427$ bits.) So, in the example above, one achieves a lower relative entropy (redundancy) by proceeding in the quantum domain, as opposed to the classical one.

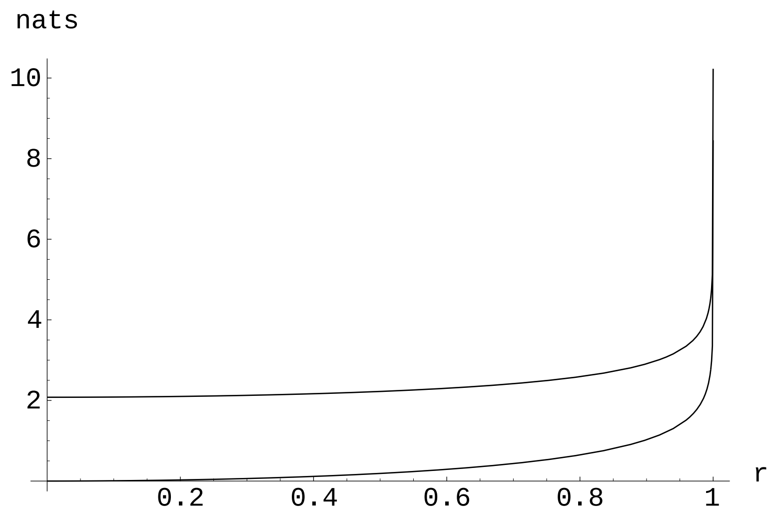

FIG. 2. Quantum asymptotic relative entropy term — $\frac{1}{2}\log I_q(r)$ — and its *larger* classical counterpart, $\frac{1}{2}\log\frac{64}{1-r^2}$, plotted against radial distance (r) in the Bloch sphere of two-level systems

In the case $r = 0$ (the fully mixed state), the quantum (Krattenthaler/Slater) asymptotics is given by the expression

$$\frac{3}{2}\log\frac{N}{2\pi \mathrm{e}} - \log w_q(0) + o(1). \tag{38}$$

For a pure state ($r = 1$), in the case that $w_q(r)$ is *continuous* and nonzero at $r = 1$, the asymptotics is given, in general, by [20]

$$2\log N - 3\log 2 - \log\pi - \log w_q(1) + o(1). \tag{39}$$

However, for the particular case of the Jeffreys' prior (31), which is *singular* at $r = 1$, we have [19, eq. (2.53)]

$$\frac{3}{2}\log N + \frac{1}{2}\log\pi - 2\log 2. \tag{40}$$

It would be of interest to ascertain if one can construct a probability distribution for which the (classical) Fisher information matrix is equal (in spherical coordinates) to [12, eq. (3.17)]

$$I_{quasi-Bures}(r,\theta,\phi) = \begin{pmatrix} \frac{1}{1-r^2} & 0 & 0 \\ 0 & \frac{r^2 g(s)}{1+r} & 0 \\ 0 & 0 & \frac{r^2 g(s)\sin^2\theta}{1+r} \end{pmatrix}, \tag{41}$$

where $s = \frac{1-r}{1+r}$ and $g(s) = \mathrm{e}s^{\frac{s}{1-s}}$. (If we employ $g(s) = \frac{2}{1+s}$ in (41), we obtain the Helstrom information matrix $H_q(r,\theta,\phi)$ [12].) This would yield the *quantum* (but non-Helstrom) information matrix, the square root of the determinant of which is proportional to the quasi-Bures probability distribution (33). This probability distribution (rather than (31), as originally conjectured [19]) has been shown to yield the common minimax and maximin in the universal coding of the two-level quantum systems [20].

C. Relations between *Monotone Metrics* and the Fisher Information Matrices Computed in Sec. III A

It would be of considerable interest to determine the precise nature $N \to \infty$ of the Fisher information matrices corresponding to the use of optimal measurements [1]. ("For the case of mixed states of spin 1/2 particles, or for higher spins we do not know what the 'outer' boundary of the set of (rescaled) achievable Fisher information matrices based on arbitrary (non separable) measurements of N systems looks like. We have some indications about the shape of this set…and we know that it is convex and compact" [2, p. 19].) In particular, we would like to ascertain whether or not there is convergence in form (to a diagonal matrix in spherical coordinates) between even and odd values of N, as numerical evidence indicates, and whether or not the Fisher information matrices are asymptotically simply proportional to some specific member (41) of a broad class of natural metric tensors (which includes the Bures and quasi-Bures metrics discussed in Sec. IV B) for the quantum states associated with operator monotone functions $f(s) = \frac{1}{g(s)}$ [12].

1. The (2,2)- and (3,3)-entries of the diagonal Fisher information matrices for even N

In fact, if we equate the (2,2)-entries of the diagonal Fisher information matrices given in sec. III B 1 for the optimal measurements for $N = 4$ and $N = 6$ to the (2,2)-cell of N times the general matrix (41) and solve for $g(s)$, recalling that $s = \frac{1-r}{1+r}$, we obtain for $N = 4$,

$$g(s) = \frac{1}{6(1+s)^3}(6 + 17s + 6s^2) \tag{42}$$

and for $N = 6$,

$$g(s) = \frac{1}{45(1+s)^5}(45 + 222s + 416s^2 + 222s^3 + 45s^4). \tag{43}$$

Both these symmetry-exhibiting functions, (42) and (43), as well as the corresponding (Bures/minimal monotone) result (the equation of a hyperbola) for $N = 2$, that is,

$$g(s) = \frac{1}{1+s} \tag{44}$$

are monotonically-decreasing on the positive real axis (Fig. 3), but we are presently not aware (for the cases $N = 4$ and 6, that is) if the reciprocals, $f(s) = 1/g(s)$, are *operator* monotone functions, as required for membership in the class of monotone metrics of Petz and Sudár [12] [39]. (A function $f(s)$, mapping the nonnegative real axis to itself, is called operator monotone if the relation $0 \leq K \leq H$ implies $0 \leq f(K) \leq f(H)$ for all matrices K and H of any order. The relation $K \leq H$ implies that all the eigenvalues of $H - K$ are nonnegative.)

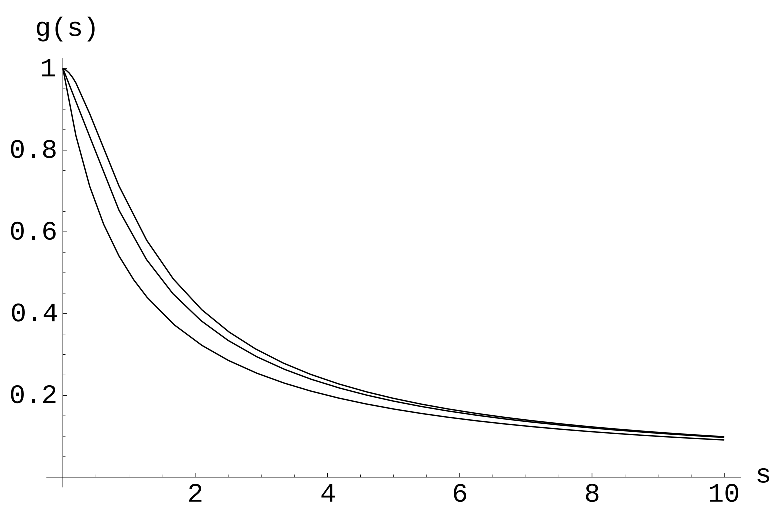

FIG. 3. Monotonically-decreasing functions $g(s)$, that is (42), (43) and (44), obtained by equating the (2,2)-entries of the computed Fisher information matrices (16), (17) and (5) for $N = 4, 6$ and 2, respectively, with N times the (2,2)-entry of the general matrix (41) for a monotone metric. The curve for $N = 6$ dominates that for $N = 4$, which in turn dominates the hyperbola for $N = 2$.

If we were to include in Fig. 3 the corresponding function for the *quasi-Bures* monotone metric, that is

$$g(s) = \frac{e s^{\frac{s}{1-s}}}{2}, \tag{45}$$

it would be essentially indistinguishable from the hyperbola for $N = 2$ (corresponding to the Bures/minimal monotone metric).

2. The (1,1)-entries of the diagonal Fisher information matrices for even N

If, pursuing these lines of thought, one could develop a formula for arbitrary (even) N for the (2,2)-entry of the Fisher information matrix for optimal measurements, and obviously easily then for the (3,3)-entry (which would be the (2,2)-entry multiplied by $\sin^2\theta$), the remaining question, of course, would be to obtain a general formula for the (1,1)-entry. In this regard, the apparent general result (established above for $N = 2, \ldots, 7$) that the Gill-Massar trace is $2N - 1$ in the pure state limit might prove helpful. But since the (1,1)-entry of the metric tensor for any monotone metric (41) is always simply $\frac{1}{1-r^2}$, it would apparently be necessary to have some *asymptotic* convergence to this expression, being that the results in the computed Fisher information matrices (16) and (17) for $N = 4$ and 6 (and presumably for arbitrary even N) contain polynomials in r in their numerators, and not simply a constant term. In Fig. 4 we plot the (1,1)-entries divided by N of the computed Fisher information matrices, in spherical coordinates, for $N = 2, 4$ and 6.

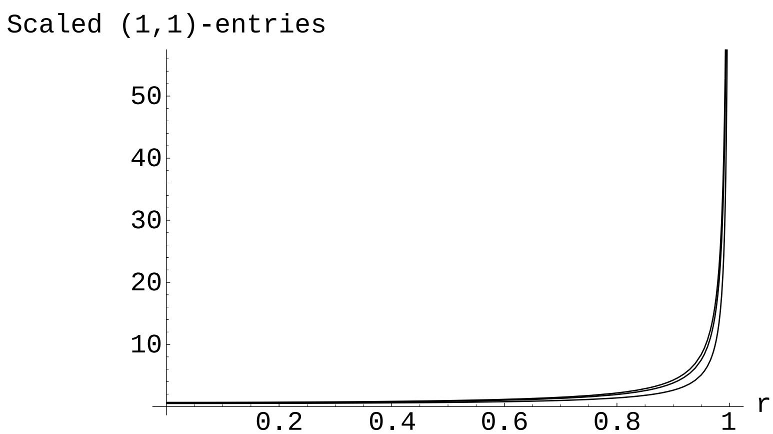

FIG. 4. (1,1)-entries divided by N of the computed diagonal Fisher information matrices (5), (16) and (17) for $N = 2, 4$ and 6, respectively. The value at $r = .9$ is greatest for $N = 6$ and least for $N = 2$.

3. Modified Gill-Massar traces based on the Yuen-Lax (maximal monotone) and quasi-Bures information matrices

In sec. III C, we defined the Gill-Massar trace as the trace of the product of the inverse of the quantum *Helstrom* information matrix and the Fisher information matrices we had computed (sec. III A) based on the optimal (in terms of *fidelity*) measurements of Vidal *et al* [1] for $N = 2, \ldots, 7$. Now the quantum Helstrom information matrix corresponds to the use of the *minimal* monotone (Bures) metric, as well as the *symmetric* logarithmic derivative. Now, we replace this with the *maximal* monotone metric, corresponding to the *right* logarithmic derivative [4, eq. (4.27)], associated with Yuen and Lax [46]. This can be accomplished by using $g(s) = (1 + s)/(2s)$ in the (diagonal/orthogonal) metric tensor (41) rather than $g(s) = \frac{2}{1+t}$ (which gives the quantum Helstrom information matrix). Then, we find that in the pure state limit ($r \to 1$) the values of the so-modified traces are exactly $N - 1$ — rather than $2N - 1$ — for all our six cases $N = 2, \ldots, 7$. For $N = 2$, this is

$$\tilde{GM}_2 = 3 - 2r^2, \tag{46}$$

for $N = 4$,

$$\tilde{GM}_4 = \frac{1}{12}(87 - 61r^2 + 10r^4), \tag{47}$$

and for $N = 6$,

$$\tilde{GM}_6 = \frac{1}{120}(1425 - 1070r^2 + 307r^4 - 62r^6). \quad (48)$$

These three functions, scaled by their value at $r = 1$, that is $N - 1$, are plotted in Fig. 5.

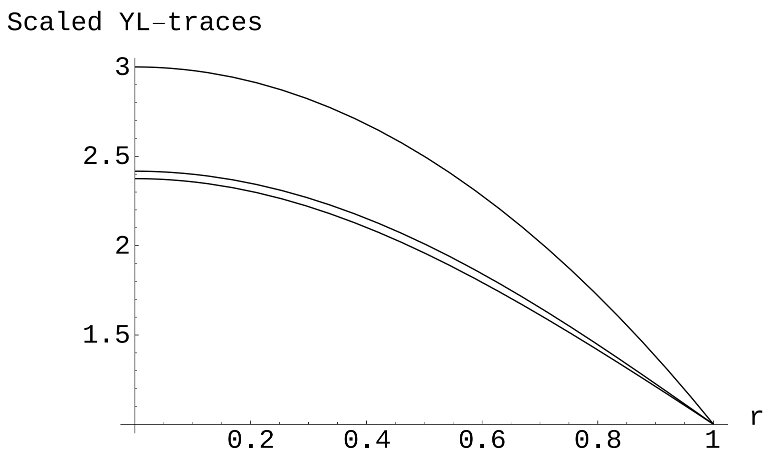

FIG. 5. Traces — scaled by $N - 1$ — for $N = 2, 4$ and 6 based on the Yuen-Lax/maximal monotone metric analysis. The y-intercepts for $r = 0$ increase with N.

The traces $\tilde{GM}_N$ for $N = 3$ and 7 are (three-line) functions of not only r, as previously, but of θ and ϕ as well. For $N = 5$, we have

$$\tilde{GM}_5 = \frac{1}{16}(147 - 96r^2 + 13r^4 + \frac{10(r^2-1)^3}{r^2 + r^2\cos 2\theta - 2}). \quad (49)$$

In the fully mixed state limit ($r \to 0$), the values of the traces are 3, 5, 7.25, 9.5, 11.875 and 11.1875.

If we alternatively employ the quasi-Bures metric, using $g(s) = es^{\frac{s}{1-s}}$, then, in the pure state limit for $N = 2, 4$ and 6 we get traces equalling $(4 + e)/e \approx 2.47152$, $3 + 8/e \approx 5.94304$ and $5 + 12/e \approx 9.41455$, respectively. (These results are intermediate, then, between those for the minimal and maximal monotone metrics.) For $r = 0$, the corresponding outcomes are the same as in the two situations above. In Fig. 6, we plot these three traces scaled by the noted values at $r = 1$.

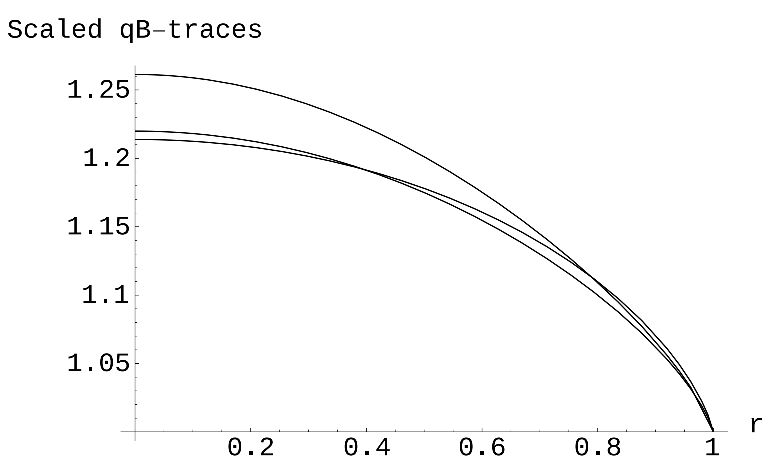

FIG. 6. Traces — scaled by their values at $r = 1$ — for $N = 2, 4$ and 6 based on the quasi-Bures monotone metric analysis. The y-intercepts for $r = 0$ increase with N.

The curves for $N = 2$ and 4 intersect at $r = .395121$.

V. CONCLUDING REMARKS

We have explicitly constructed the 3×3 Fisher information matrices for the optimal measurements of Vidal *et al* [1] for $N = 2, \ldots, 7$, found that they are tightly bounded by $(N-1)H_q$ near the pure state boundary, and conjectured that they converge from above to $\frac{N}{2}$ times the identity matrix at the fully mixed state ($r = 0$). As our main finding, we have uncovered (sec. III C) an interesting (less strict) analogue for non-separable measurements of a "new quantum Cramér-Rao inequality" of Gill and Massar [2, eq. (27)]. The possibility of extending it to the cases $N > 7$ appears to be a challenging problem. Also, the development of optimal measurement schemes for multiple copies of m-level systems, $m > 2$, and the subsequent evaluation of their Fisher information characteristics, merits investigation (cf. [9]). In this regard, we have presented in sec. III D 3 additional evidence — for an optimal measurement we devised for the case $m = 3$, $N = 2$ — that has led us to the conjecture that for optimal non-separable measurements of N copies of m-level quantum systems, the "Gill-Massar trace" equals $(2N-1)(m-1)$ in the pure state limit for *all* m and N.

Additionally, it would be of interest to study the Fisher information matrices associated with optimal measurements based on *continuous* oproms [47, p. 386] [48]. The relation between optimal measurements (sec. III) and universal quantum coding (sec. IV B)— both involving averaging with respect to isotropic prior probability distributions by projecting onto total spin eigenstates — appears to be worthy of further consideration. (Fischer and Freyberger recently compared the use of single adaptive measurements — which possess certain practical advantages — with the use of non-separable ones [3].)

We have also investigated here several related topics, all pertaining to the information-theoretic properties of the two-level quantum systems. We have posed the problem of constructing an operator-valued probability measure (oprom) for the smallest number possible of copies $N \geq 4$ which yields the quadrinomial probability distribution (2), the Fisher information matrix for which is simply four times the quantum (Helstrom) information matrix (2). Also, we discuss in sec. III A 6 what appears to be an intriguing connection between our results and the work of Frieden [29] concerning differences between classical and quantum information.

ACKNOWLEDGMENTS

I would like to express appreciation to the Institute for Theoretical Physics for computational support in this research, as well as to M. J. W. Hall, G. Vidal, R. Tarrach, R. Gill and B. R. Frieden for various forms of assistance and advice.

---

[1] G. Vidal, J. I. Latorre, P. Pascual, and R. Tarrach, Phys. Rev. A 60, 126 (1999).
[2] R. D. Gill and S. Massar, Phys. Rev. A 61, 042312/1-16 (2000).
[3] D. G. Fischer and M. Freyberger, *Estimating Mixed Quantum States*, quant-ph/0005090.
[4] C. W. Helstrom, *Quantum Detection and Estimation Theory*, (Academic, New York, 1976).
[5] R. Gill, *Asymptotics in Quantum Statistics*, (Mathematical Institute, University of Utrecht, 1999). available at WWW: http://math.uu.nl/people/gill/Preprints/paper.ps.gz.
[6] P. Busch, G. Cassinelli, and P. J. Lahti, Revs. Math. Phys. 7, 1105 (1995).
[7] R. Tarrach and G. Vidal, Phys. Rev. A 60, R3339 (1999).
[8] J. I. Latorre, P. Pascual, and R. Tarrach, Phys. Rev. Lett. 81, 1351 (1998).
[9] A. Acín, J. I. Latorre, and P. Pascual, Phys. Rev. A 61, 022113/1-7 (2000).
[10] A. Uhlmann, Rep. Math. Phys. 9, 273 (1976).
[11] R. Jozsa, J. Mod. Opt. 41, 2315 (1994).
[12] D. Petz and C. Sudár, J. Math. Phys. 37, 2662 (1996).
[13] S. L. Braunstein and C. M. Caves, Phys. Rev. Lett. 72, 3439 (1994).
[14] P. B. Slater, Phys. Lett. A 247, 1 (1998).
[15] S. L. Braunstein and G. J. Milburn, Phys. Rev. A 51, 1820 (1995).
[16] E. G. Beltrametti and G. Cassinelli, *The Logic of Quantum Mechanics*, (Addison-Wesley, Reading, 1981).
[17] P. B. Slater, J. Math. Phys. 37, 2682 (1996).
[18] B. S. Clarke and A. R. Barron, IEEE Info. Th. 36, 453 (1990).
[19] C. Krattenthaler and P. B. Slater, Trans. IEEE Info. Th. 46, 801 (2000).

[20] H. Grosse, C. Krattenthaler, and P. B. Slater, *Asymptotic Redundancies for Universal Quantum Coding. II* (in preparation).
[21] R. Jozsa, M. Horodecki, P. Horodecki, and R. Horodecki, Phys. Rev. Lett. 81, 1714 (1998).
[22] J. Dittmann, Sem. Sophus Lie, 3, 73 (1993).
[23] J. Dittmann, J. Phys. A 32, 2663 (1999).
[24] O. E. Barndorff-Nielsen and R. D. Gill, J. Phys. A 33, 4481 (2000).
[25] M. Hübner, Phys. Lett. A 163, 239 (1992).
[26] M. Hübner, Phys. Lett. A 179, 226 (1993).
[27] A. Fujiwara and H. Nagaoka, Phys. Lett. A 201, 119 (1995).
[28] K. P. Tod, Class. Quant. Grav. 9, 1693 (1992).
[29] B. R. Frieden, *Physics from Fisher Information: A Unification*, (Cambridge University Press, Cambridge, 1999).
[30] M. K. Murray and J. W. Rice, *Differential Geometry and Statistics*, (Chapman and Hall, London, 1993).
[31] R. E. Kass, Statist. Sci. 4, 188 (1989).
[32] A. M. Kagan, Probl. Pered. Inform. 12(2), 20 (1976).
[33] N. N. Chentsov, in *Encyclopaedia of Mathematics*, edited by M. Hazewinkel (Kluwer, Dordrecht, 1990), vol. 5, p. 78.
[34] A. M. Kagan and Z. Landsman, Stat. Prob. Lett. 32, 175 (1997).
[35] C. R. Rao, *Linear Statistical Inference and Its Applications* (Wiley, New York, 1973).
[36] C. H. Bennett, D. P. DiVincenzo, C. A. Fuchs, T. Mor, E. Rains, P. W. Shor, J. A. Smolin, and W. K. Wootters, Phys. Rev. A 59, 1070 (1999).
[37] D. R. Cox and N. Reid, J. R. Statist. Soc. B 49, 1 (1987).
[38] S. Kobayashi and K. Nomizu, *Foundations of Differential Geometry. Vol. 1*, (Interscience, New York, 1963).
[39] A. Lesniewski and M. B. Ruskai, J. Math. Phys. 40, 5702 (1999).
[40] A. Fujiwara and H. Nagaoka, J. Math. Phys. 40, 4227 (1999).
[41] C. M. Caves and G. J. Milburn, Opt. Commun. 179, 439 (2000).
[42] P. B. Slater, *Bures Geometry of the Three-Level Quantum Systems*, quant-ph/0008069.
[43] M. S. Byrd and P. B. Slater, *Bures Measures over the Spaces of Two and Three-Dimensional Density Matrices*, quant-ph/0004055 (to appear in Phys. Lett. A).
[44] V. E. Mkrtchian and V. O. Chaltykian, Opt. Commun. 63, 239 (1987).
[45] P. B. Slater, J. Phys. A 32, 8231 (1999).
[46] H. P. Yuen and M. Lax, Trans. IEEE Info. Th. 19, 740 (1973).
[47] A. Peres, *Quantum Theory: Concepts and Methods*, (Kluwer, Dordrecht, 1995).
[48] P. B. Slater, J. Math. Phys. 38, 2274 (1997).

List of Figures

1 Gill-Massar traces for $N = 4, 5, 6$ and 7 scaled by their values at the pure states, $r = 1$, that is, $2N - 1$. The y-intercepts for $r = 0$, corresponding to the fully mixed state, increase with N. 8
2 Quantum asymptotic relative entropy term — $\frac{1}{2}\log I_q(r)$ — and its *larger* classical counterpart, $\frac{1}{2}\log\frac{64}{1-r^2}$, plotted against radial distance (r) in the Bloch sphere of two-level systems 11
3 Monotonically-decreasing functions $g(s)$, that is (42), (43) and (44), obtained by equating the (2,2)-entries of the computed Fisher information matrices (16), (17) and (5) for $N = 4, 6$ and 2, respectively, with N times the (2,2)-entry of the general matrix (41) for a monotone metric. The curve for $N = 6$ dominates that for $N = 4$, which in turn dominates the hyperbola for $N = 2$. 12
4 (1,1)-entries divided by N of the computed diagonal Fisher information matrices (5), (16) and (17) for $N = 2, 4$ and 6, respectively. The value at $r = .9$ is greatest for $N = 6$ and least for $N = 2$. 13
5 Traces — scaled by $N - 1$ — for $N = 2, 4$ and 6 based on the Yuen-Lax/maximal monotone metric analysis. The y-intercepts for $r = 0$ increase with N. 14
6 Traces — scaled by their values at $r = 1$ — for $N = 2, 4$ and 6 based on the quasi-Bures monotone metric analysis. The y-intercepts for $r = 0$ increase with N. 14